



PCT

## RAW SEQUENCE LISTING

DATE: 10/08/2004

PATENT APPLICATION: US/10/509,785

TIME: 11:58:21

Input Set : A:\3462.1010-000.txt

Output Set: N:\CRF4\10082004\J509785.raw

3 <110> APPLICANT: Narimatsu, Hisashi  
 4 Kudo, Takashi  
 5 Iwasaki, Hiroko  
 7 <120> TITLE OF INVENTION: NOVEL GALACTOSYLTRANSFERASES, THEIR PEPTIDES, AND  
 8 NUCLEIC ACIDS ENCODING THE SAME  
 10 <130> FILE REFERENCE: 3462.1010-000  
 C--> 12 <140> CURRENT APPLICATION NUMBER: US/10/509,785  
 C--> 13 <141> CURRENT FILING DATE: 2004-09-29  
 15 <150> PRIOR APPLICATION NUMBER: JP 2002-94772  
 16 <151> PRIOR FILING DATE: 2002-03-29  
 18 <150> PRIOR APPLICATION NUMBER: JP 2002-201344  
 19 <151> PRIOR FILING DATE: 2002-07-10  
 21 <160> NUMBER OF SEQ ID NOS: 24  
 23 <170> SOFTWARE: PatentIn Ver. 2.0  
 25 <210> SEQ ID NO: 1  
 26 <211> LENGTH: 1471  
 27 <212> TYPE: DNA  
 28 <213> ORGANISM: Homo sapiens  
 30 <220> FEATURE:  
 31 <221> NAME/KEY: CDS  
 32 <222> LOCATION: (105)..(1058)  
 34 <400> SEQUENCE: 1  
 35 agaacagcct ggtcaggagc gtaacggagt ggtgcgccaa cgtgagagga aaccctgtgcg 60  
 37 cggtcgcgt ttcctgtccc caagccgttc tagacgcggg aaaa atg ctt tct gaa 116  
 38 Met Leu Ser Glu  
 39 1  
 41 agc agc tcc ttt ttg aag ggt gtg atg ctt gga agc att ttc tgt gct 164  
 42 Ser Ser Ser Phe Leu Lys Gly Val Met Leu Gly Ser Ile Phe Cys Ala  
 43 5 10 15 20  
 45 ttg atc act atg cta gga cac att agg att ggt cat gga aat aga atg 212  
 46 Leu Ile Thr Met Leu Gly His Ile Arg Ile Gly His Gly Asn Arg Met  
 47 25 30 35  
 49 cac cac cat gag cat cat cac cta caa gct cct aac aaa gaa gat atc 260  
 50 His His His Glu His His His Leu Gln Ala Pro Asn Lys Glu Asp Ile  
 51 40 45 50  
 53 ttg aaa att tca gag gat gag cgc atg gag ctc agt aag agc ttt cga 308  
 54 Leu Lys Ile Ser Glu Asp Glu Arg Met Glu Leu Ser Lys Ser Phe Arg  
 55 55 60 65  
 57 gta tac tgt att atc ctt gta aaa ccc aaa gat gtg agt ctt tgg gct 356  
 58 Val Tyr Cys Ile Ile Leu Val Lys Pro Lys Asp Val Ser Leu Trp Ala  
 59 70 75 80  
 61 gca gta aag gag act tgg acc aaa cac tgt gac aaa gca gag ttc ttc 404  
 62 Ala Val Lys Glu Thr Trp Thr Lys His Cys Asp Lys Ala Glu Phe Phe

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/509,785

DATE: 10/08/2004

TIME: 11:58:21

Input Set : A:\3462.1010-000.txt

Output Set: N:\CRF4\10082004\J509785.raw

```

63 85          90          95          100
65 agt tct gaa aat gtt aaa gtg ttt gag tca att aat atg gac aca aat 452
66 Ser Ser Glu Asn Val Lys Val Phe Glu Ser Ile Asn Met Asp Thr Asn
67          105          110          115
69 gac atg tgg tta atg atg aga aaa gct tac aaa tac gcc ttt gat aag 500
70 Asp Met Trp Leu Met Met Arg Lys Ala Tyr Lys Tyr Ala Phe Asp Lys
71          120          125          130
73 tat aga gac caa tac aac tgg ttc ttc ctt gca cgc ccc act acg ttt 548
74 Tyr Arg Asp Gln Tyr Asn Trp Phe Phe Leu Ala Arg Pro Thr Thr Phe
75          135          140          145
77 gct atc att gaa aac cta aag tat ttt ttg tta aaa aag gat cca tca 596
78 Ala Ile Ile Glu Asn Leu Lys Tyr Phe Leu Leu Lys Lys Asp Pro Ser
79          150          155          160
81 cag cct ttc tat cta ggc cac act ata aaa tct gga gac ctt gaa tat 644
82 Gln Pro Phe Tyr Leu Gly His Thr Ile Lys Ser Gly Asp Leu Glu Tyr
83 165          170          175          180
85 gtg ggt atg gaa gga gga att gtc tta agt gta gaa tca atg aaa aga 692
86 Val Gly Met Glu Gly Gly Ile Val Leu Ser Val Glu Ser Met Lys Arg
87          185          190          195
89 ctt aac agc ctt ctc aat atc cca gaa aag tgt cct gaa cag gga ggg 740
90 Leu Asn Ser Leu Leu Asn Ile Pro Glu Lys Cys Pro Glu Gln Gly Gly
91          200          205          210
93 atg att tgg aag ata tct gaa gat aaa cag cta gca gtt tgc ctg aaa 788
94 Met Ile Trp Lys Ile Ser Glu Asp Lys Gln Leu Ala Val Cys Leu Lys
95          215          220          225
97 tat gct gga gta ttt gca gaa aat gca gaa gat gct gat gga aaa gat 836
98 Tyr Ala Gly Val Phe Ala Glu Asn Ala Glu Asp Ala Asp Gly Lys Asp
99          230          235          240
101 gta ttt aat acc aaa tct gtt ggg ctt tct att aaa gag gca atg act 884
102 Val Phe Asn Thr Lys Ser Val Gly Leu Ser Ile Lys Glu Ala Met Thr
103 245          250          255          260
105 tat cac ccc aac cag gta gta gaa ggc tgt tgt tca gat atg gct gtt 932
106 Tyr His Pro Asn Gln Val Val Glu Gly Cys Cys Ser Asp Met Ala Val
107          265          270          275
109 act ttt aat gga ctg act cca aat cag atg cat gtg atg atg tat ggg 980
110 Thr Phe Asn Gly Leu Thr Pro Asn Gln Met His Val Met Met Tyr Gly
111          280          285          290
113 gta tac cgc ctt agg gca ttt ggg cat att ttc aat gat gca ttg gtt 1028
114 Val Tyr Arg Leu Arg Ala Phe Gly His Ile Phe Asn Asp Ala Leu Val
115          295          300          305
117 ttc tta cct cca aat ggt tct gac aat gac tgagaagtgg tagaaaagcg 1078
118 Phe Leu Pro Pro Asn Gly Ser Asp Asn Asp
119          310          315
121 tgaatatgat ctttgtatag gacgtgtggt gtcattatatt gtagtagtaa ctacatatcc 1138
123 aatacagctg tatgtttctt tttcttttct aatttggtgg cactgggtata accacacatt 1198
125 aaagtacagta gtacatTTTT aaatgagggt ggTTTTTTTT tttaaaacac atgaacattg 1258
127 taaatgtggt ggaaagaagt gttttaagaa taataatttt gcaataaatac tattaataaa 1318
129 tattatatgt gataaattct aaattatgaa cattagaaat ctgtggggca catatttttg 1378
131 ctgattgggt aaaaaatttt aacaggtctt tagcgttcta agatatgcaa atgatatctc 1438

```

## RAW SEQUENCE LISTING

DATE: 10/08/2004

PATENT APPLICATION: US/10/509,785

TIME: 11:58:21

Input Set : A:\3462.1010-000.txt

Output Set: N:\CRF4\10082004\J509785.raw

```

133 tagttgtgaa tttgtgatta aagtaaaact ttt                                     1471
136 <210> SEQ ID NO: 2
137 <211> LENGTH: 318
138 <212> TYPE: PRT
139 <213> ORGANISM: Homo sapiens
141 <400> SEQUENCE: 2
142 Met Leu Ser Glu Ser Ser Ser Phe Leu Lys Gly Val Met Leu Gly Ser
143   1           5           10           15
145 Ile Phe Cys Ala Leu Ile Thr Met Leu Gly His Ile Arg Ile Gly His
146           20           25           30
148 Gly Asn Arg Met His His His Glu His His His Leu Gln Ala Pro Asn
149           35           40           45
151 Lys Glu Asp Ile Leu Lys Ile Ser Glu Asp Glu Arg Met Glu Leu Ser
152           50           55           60
154 Lys Ser Phe Arg Val Tyr Cys Ile Ile Leu Val Lys Pro Lys Asp Val
155           65           70           75           80
157 Ser Leu Trp Ala Ala Val Lys Glu Thr Trp Thr Lys His Cys Asp Lys
158           85           90           95
160 Ala Glu Phe Phe Ser Ser Glu Asn Val Lys Val Phe Glu Ser Ile Asn
161           100          105          110
163 Met Asp Thr Asn Asp Met Trp Leu Met Met Arg Lys Ala Tyr Lys Tyr
164           115          120          125
166 Ala Phe Asp Lys Tyr Arg Asp Gln Tyr Asn Trp Phe Phe Leu Ala Arg
167           130          135          140
169 Pro Thr Thr Phe Ala Ile Ile Glu Asn Leu Lys Tyr Phe Leu Leu Lys
170 145          150          155          160
172 Lys Asp Pro Ser Gln Pro Phe Tyr Leu Gly His Thr Ile Lys Ser Gly
173          165          170          175
175 Asp Leu Glu Tyr Val Gly Met Glu Gly Gly Ile Val Leu Ser Val Glu
176          180          185          190
178 Ser Met Lys Arg Leu Asn Ser Leu Leu Asn Ile Pro Glu Lys Cys Pro
179          195          200          205
181 Glu Gln Gly Gly Met Ile Trp Lys Ile Ser Glu Asp Lys Gln Leu Ala
182          210          215          220
184 Val Cys Leu Lys Tyr Ala Gly Val Phe Ala Glu Asn Ala Glu Asp Ala
185 225          230          235          240
187 Asp Gly Lys Asp Val Phe Asn Thr Lys Ser Val Gly Leu Ser Ile Lys
188          245          250          255
190 Glu Ala Met Thr Tyr His Pro Asn Gln Val Val Glu Gly Cys Cys Ser
191          260          265          270
193 Asp Met Ala Val Thr Phe Asn Gly Leu Thr Pro Asn Gln Met His Val
194          275          280          285
196 Met Met Tyr Gly Val Tyr Arg Leu Arg Ala Phe Gly His Ile Phe Asn
197          290          295          300
199 Asp Ala Leu Val Phe Leu Pro Pro Asn Gly Ser Asp Asn Asp
200 305          310          315
203 <210> SEQ ID NO: 3
204 <211> LENGTH: 30
205 <212> TYPE: DNA

```

## RAW SEQUENCE LISTING

DATE: 10/08/2004

PATENT APPLICATION: US/10/509,785

TIME: 11:58:21

Input Set : A:\3462.1010-000.txt

Output Set: N:\CRF4\10082004\J509785.raw

```

206 <213> ORGANISM: Artificial Sequence
208 <220> FEATURE:
209 <223> OTHER INFORMATION: Description of Artificial Sequence:Artificially
210     Synthesized Primer Sequence
212 <400> SEQUENCE: 3
213 gaagatctag aatgcaccac catgagcatc          30
216 <210> SEQ ID NO: 4
217 <211> LENGTH: 39
218 <212> TYPE: DNA
219 <213> ORGANISM: Artificial Sequence
221 <220> FEATURE:
222 <223> OTHER INFORMATION: Description of Artificial Sequence:Artificially
223     Synthesized Primer Sequence
225 <400> SEQUENCE: 4
226 ataagaatgc ggccgctcag tcattgtcag aaccatttg          39
229 <210> SEQ ID NO: 5
230 <211> LENGTH: 67
231 <212> TYPE: DNA
232 <213> ORGANISM: Artificial Sequence
234 <220> FEATURE:
235 <223> OTHER INFORMATION: Description of Artificial Sequence:Artificially
236     Synthesized Primer Sequence
238 <400> SEQUENCE: 5
239 ggggacaagt ttgtacaaaa aagcaggctt agaaggagat agaaccatgc tttctgaaag 60
241 cagctcc          67
244 <210> SEQ ID NO: 6
245 <211> LENGTH: 50
246 <212> TYPE: DNA
247 <213> ORGANISM: Artificial Sequence
249 <220> FEATURE:
250 <223> OTHER INFORMATION: Description of Artificial Sequence:Artificially
251     Synthesized Primer Sequence
253 <400> SEQUENCE: 6
254 ggggaccact ttgtacaaga aagctgggtc tcaatcattg tcagaaccat          50
257 <210> SEQ ID NO: 7
258 <211> LENGTH: 19
259 <212> TYPE: PRT
260 <213> ORGANISM: Artificial Sequence
262 <220> FEATURE:
263 <223> OTHER INFORMATION: Description of Artificial Sequence:Artificially
264     Synthesized Peptide Sequence
266 <400> SEQUENCE: 7
267 Val Pro Ser Thr Pro Pro Thr Pro Ser Pro Ser Thr Pro Pro Thr Pro
268   1           5           10           15
270 Ser Pro Ser
273 <210> SEQ ID NO: 8
274 <211> LENGTH: 25
275 <212> TYPE: DNA
276 <213> ORGANISM: Artificial Sequence

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/509,785

DATE: 10/08/2004

TIME: 11:58:21

Input Set : A:\3462.1010-000.txt

Output Set: N:\CRF4\10082004\J509785.raw

```

278 <220> FEATURE:
279 <223> OTHER INFORMATION: Description of Artificial Sequence:Artificially
280     Synthesized Primer Sequence
282 <400> SEQUENCE: 8
283 gtttgctga aatatgctgg agtat                25
286 <210> SEQ ID NO: 9
287 <211> LENGTH: 24
288 <212> TYPE: DNA
289 <213> ORGANISM: Artificial Sequence
291 <220> FEATURE:
292 <223> OTHER INFORMATION: Description of Artificial Sequence:Artificially
293     Synthesized Primer Sequence
295 <400> SEQUENCE: 9
296 caacagcctt ctactacctg gttg                24
299 <210> SEQ ID NO: 10
300 <211> LENGTH: 35
301 <212> TYPE: DNA
302 <213> ORGANISM: Artificial Sequence
304 <220> FEATURE:
305 <223> OTHER INFORMATION: Description of Artificial Sequence:Artificially
306     Synthesized Probe Sequence
308 <400> SEQUENCE: 10
309 cagaaaatgc agaagatgct gatggaaaag atgta    35
312 <210> SEQ ID NO: 11
313 <211> LENGTH: 19
314 <212> TYPE: DNA
315 <213> ORGANISM: Artificial Sequence
317 <220> FEATURE:
318 <223> OTHER INFORMATION: Description of Artificial Sequence:Artificially
319     Synthesized Primer Sequence
321 <400> SEQUENCE: 11
322 agaaatacac ttctcgggaa                    19
325 <210> SEQ ID NO: 12
326 <211> LENGTH: 20
327 <212> TYPE: DNA
328 <213> ORGANISM: Artificial Sequence
330 <220> FEATURE:
331 <223> OTHER INFORMATION: Description of Artificial Sequence:Artificially
332     Synthesized Primer Sequence
334 <400> SEQUENCE: 12
335 tgcagtgcta gacatattac                    20
338 <210> SEQ ID NO: 13
339 <211> LENGTH: 23
340 <212> TYPE: DNA
341 <213> ORGANISM: Artificial Sequence
343 <220> FEATURE:
344 <223> OTHER INFORMATION: Description of Artificial Sequence:Artificially
345     Synthesized Primer Sequence
347 <400> SEQUENCE: 13

```

**VERIFICATION SUMMARY**

PATENT APPLICATION: US/10/509,785

DATE: 10/08/2004

TIME: 11:58:22

Input Set : A:\3462.1010-000.txt

Output Set: N:\CRF4\10082004\J509785.raw

L:12 M:270 C: Current Application Number differs, Replaced Current Application Number

L:13 M:271 C: Current Filing Date differs, Replaced Current Filing Date